33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Integrating the spatial profile of the N200 speller for asynchronous brain-computer interfaces

Dan Zhang, Honglai Xu, Wei Wu, Shangkai Gao, *Fellow IEEE*, Bo Hong*, *Member, IEEE*

***Abstract*—The N200 speller is a novel brain-computer interface (BCI) paradigm utilizing the overt attention effects on motion onset visual evoked potentials (mVEP). However, the asynchronous performance of the N200 BCI has not been fully explored. In this paper, a novel algorithm was proposed, integrating the spatial profile of the visual speller to provide a more precise description of the mVEP responses. Most importantly, only control state data were used in the algorithm to train a classifier which can detect the non-control state effectively. Using offline recorded data, the asynchronous performance of the proposed algorithm was shown to be significantly better than that of a similar algorithm without using the spatial information. The proposed algorithm can be used for developing a practical, asynchronous N200 BCI system.**

I. Introduction

The N200 speller is a recently developed brain-computer interface system, using motion-onset visual evoked potentials (mVEPs) as neural signals for translating user intention into device commands [1-2]. Similar to the visual speller used in the classical P300 speller proposed by Farwell and Donchin [3-4], a 6×6 virtual button matrix is employed in the N200 speller BCI for its high efficiency of locating the target virtual button. Instead of using flashing rows and columns for the target-related P300 component, brief motions of chromatic visual objects were embedded in the 36 virtual buttons to evoke a motion-onset specific N200 component, which is the most prominent component of mVEPs. In the N200 speller paradigm, the users selected one virtual button by overtly attending to (i.e. gazing at) the chromatic moving bars appearing at the spatial location of the button. The motion stimuli corresponding to the target button elicited distinct mVEPs different from those of the non-target (unattended) buttons, which was the basis for BCI classification.

mVEPs were believed to be generated from the extrastriate temporo-occipital and associated parietal cortical areas [5-6]. Compared with other types of visual evoked potentials (VEP), mVEPs have been shown to have low inter- and intra-subject variability [5, 7] and visual stimuli with relatively low contrast and luminance level is sufficient to evoke stable responses [5, 8-9]. Hereby, mVEP is a promising candidate for building a practical BCI system since the above-mentioned advantages can result in less subject-specific calibration time, longer system stability, greater convenience to be integrated into other applications and less user fatigue. To date, the N200 speller has been reported to achieve a comparable target detection accuracy with that of the P300-speller [2] and an online N200 BCI integrated in a Google search system with 6 virtual buttons has been implemented [10].

However, previous reported N200 BCIs worked in a synchronous way, i.e. the user was always assumed to be in a control state. This can be problematic when the BCI system is used in daily situation: the users need the system to stay silent when they are resting, watching TV etc. Therefore, the non-control state has to be taken into account for developing a practical N200 BCI. In the past decades, numerous efforts have been made to develop algorithms for asynchronous P300 BCIs [13-15]. Although some of the algorithms for P300 BCIs can be applied in N200 BCIs as well, they didn't fully explore the characteristics of mVEPs. To our knowledge, no such studies specifically for N200 BCIs have been reported.

One important feature of mVEPs is that they are recordable up to about 50° in the periphery of the visual field [11] and they have a much lower amplitude decrease with retinal eccentricity compared with pattern-reversal VEPs [12]. Hence, the motion stimuli from the non-target (unattended) buttons in the 6×6 virtual button matrix will also elicit mVEPs but the amplitude of these responses will decrease with the physical distance between these buttons and the target button becoming larger. However, for most algorithms for P300 BCIs [13-15] as well as the previous algorithms used for N200 BCIs [1, 10], all the non-target responses were regarded as one class for classification, regardless of their physical distance to the target. Nevertheless, the spatial profile may provide useful information for BCI classification: the mVEPs evoked by the virtual button next to the target button are certainly larger in amplitude than the ones evoked by the virtual button further way from the target. By integrating this spatial profile of the N200 speller, better recognition of the control state is expected. More importantly, the non-control state can be detected better since having mVEPs responses fit for the spatial profile is a more strict criterion than simply judging whether one particular mVEP response is target or not. Another advantage by introducing the spatial profile is this new classifier is supposed to discriminate the control and non-control state data based on

Manuscript received March 30, 2011. This work was supported in part by the National Science Foundation of China under Grant 61071003 and 90820304.

B. Hong is with the Department of Biomedical Engineering, Tsinghua University, Beijing 100084, China (e-mai: hongbo@tsinghua.edu.cn).

D. Zhang, H. Xu, W. Wu and S. Gao are with Department of Biomedical Engineering, Tsinghua University, Beijing 100084, China.

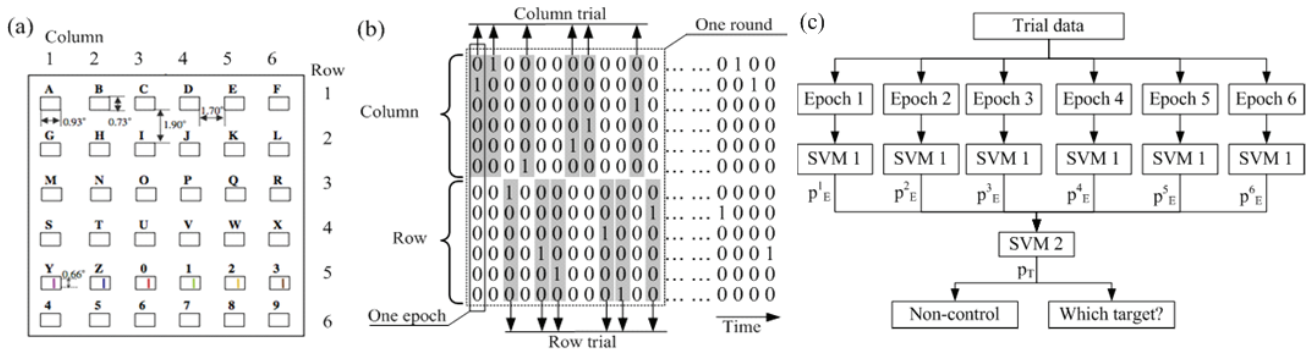

Fig. 1. (a) the N200 speller interface, note the index (1-6) for rows and columns; (b) Stimulus sequence of the N200 speller, where 1 denotes the presentation of motion stimuli from certain column/row and 0 denotes no stimuli. Epoch, trial, and round are defined in (b); (c) Procedure of the proposed algorithm.

how well the data fit the responses according to the spatial profile, thus no non-control state data were needed for training the classifier.

In this paper, a two-layer algorithm considering the spatial profile of the N200 speller was proposed. A two-class classifier was constructed as the first layer to encode the spatial profile information in a feature space and a second six-class classifier was employed to identify which row/column the subject was attending to. The recognition of the non-control state was achieved by applying a probability-based threshold in the second classifier. Compared with a traditional classifier not using the spatial information, the proposed algorithm showed better classification accuracy for both the control and non-control state, on the same dataset.

II. Methods

A. Data description

An illustration of the N200 visual speller interface was shown in Fig. 1a. In each virtual button, a vertical bar with a height of 0.66° visual angle appeared (motion-onset) at the right border of a vacant rectangle and moved leftward at the velocity of 3.10°/s before it disappeared (motion offset), forming a brief motion stimulus. The entire process of onset, motion and offset took 140 ms. The stimulus onset asynchrony (SOA) between two motion stimuli was 200 ms. The motion stimuli in the virtual buttons appeared by row/column, with random color. One round of stimulus presentation consisted of 12 stimuli in a random sequence, corresponding to the six rows and six columns respectively.

Ten healthy subjects (six male and four female, aged 20–28) participated in the experiment. Within each experiment block, they were instructed to overtly attend to one virtual button as the target for 15 rounds. All subjects completed two experiment sessions in an offline manner, each consisting of 6 blocks with target buttons along the diagonal of the speller matrix from the upper left to the lower right corner. EEG data were acquired from 30 surface electrodes using a NeuroScan SynAmp II amplifier at the sampling rate of 200Hz. Please refer to Hong et al. [2] for more details on the experiment paradigm and the data acquisition.

Selection of features such as channels, time windows was beyond the scope of this paper since we focused on the classification algorithm. Therefore, we used the same procedure as reported in the previous studies [2, 10, 16]: the EEG data were first segmented according to stimulus onset and downsampled to 20Hz. Data from channels P3, P7, O1 between 100 and 500ms following the stimulus onset were used, resulting in a 27-dimension feature vector per stimulus. The selected channels have been reported previously to exhibit strong mVEP responses [2].

B. Problem formulation

The data from the N200 speller are organized in an *epoch-trial-block* structure. Here an *epoch* is associated with one particular motion stimulus (presented in either row or column) and represents the EEG data in the corresponding time segment. The *epochs* are categorized into *target epochs* and *non-target epochs*, depending on the subject's attention task. Within each round of stimulus presentation, the EEG data can be separated into a *column trial* and a *row trial*, each including 6 epochs, as depicted in Fig. 1b. By identifying the target epochs from both the column and row trials from the same round, the target virtual button can be located. 15 continuously recorded column/row trials with the same attention task are defined as one *block*. Column trials and row trials were not discriminated in the proposed algorithm since they showed similar characteristics [2] and shared the same spatial profile. In the dataset described in II.A, each subject's data consist of 12 blocks, 180 trials, and 2160 epochs. 1/6 of the epochs are target epochs.

To exploit the spatial profile, a variable d in units of the distance between two nearby rows/columns is introduced to measure the physical distance between two rows or columns, i.e. the physical distance between column 2 and 6 is 4. Thus, the non-target epochs can be categorized into 5 types according to their d values.

Considering the nice eccentricity effects of mVEPs [11], we hypothesize that the 6 epochs from the same trial will preserve a certain response structure when the user is in the control state, showing the largest mVEP amplitude (i.e. N200) for the target row/column and a decrease of the mVEPs with increasing physical distances (d) to the target. This response

structure is the *spatial profile* we seek to learn from the data. Accordingly, there are 6 trial types, each with distinct spatial profiles based on the series of *d*-values associated with their attention tasks, e.g. trials with row/column 1 as target contain the epochs with *d*={0, 1, 2, 3, 4, 5} (from left to right of the N200 speller interface) while trials with column 3 as target have epochs with *d*={2, 1, 0, 1, 2, 3} Trials without such a recognizable structure will be regarded as in the non-control state.

With such a criterion, a more precise description of the brain status associated with the N200 speller was achieved. The procedure of the proposed algorithm was shown in Fig. 1c. The details were described in the following sections.

C. Encoding of the spatial profile in a feature space

Overt attention resulted in larger amplitude responses of the N200 component of mVEPs. However, simply using the N200 component as the feature for classification might be problematic due to the intra- and inter-subject variability. To avoid such problems, a two-class classifier using support vector machine (SVM) was constructed to project the information from the spatial profile into a feature space. In contrast to the traditional methods where all the non-target epochs were used for training the classifier, we used only the features (27-dim vector as described in 2.1) from the non-target epochs with *d*=4/5 together with target epochs (*d*=0) as the input to the SVM classifier. For each testing epoch X_E, the output of the classifier was further translated into (epoch-based) p-value (p_E), describing the probability of one particular epoch to be a target epoch.

$$p_E = SVM_{Layer-1}(X_E) \tag{1}$$

By constructing the training dataset in this way, the epochs with *d*=1, 2, 3 (not used in training dataset) were expected to yield decreasing p-values with increasing *d*, preserving the spatial profile information.

D. Trial-based classification

After performing II.C, the EEG data from one trial were transformed into a 6-dimension feature vector, representing the p-values of the 6 epochs. A 6-class SVM classifier was employed for the classification of the 6 types of trials. Again, we obtain the probability output from the SVM classifier, resulting in 6 (trial-based) p-values per trial ($p_T^k, k = 1,...,6$), indicating how likely the given trial belonged to one trial type. One trial will be classified as in the non-control state when all the 6 trial-based p-values are below certain pre-defined threshold (p_{thr}).

$$\{p_T^k\}_{k=1}^6 = SVM_{Layer-2}(\{p_E^i\}_{i=1}^6) \tag{2}$$

$$O_T = \begin{cases} 0, if\ \max(p_T^k) < p_{thr} \\ K, if\ \max(p_T^k) = p_T^K, p_T^K > p_{thr} \end{cases} \tag{3}$$

E. Evaluation of the proposed algorithm

To evaluate the performance of the proposed algorithm, a second algorithm following the same structure but without utilizing the spatial profile was implemented for comparison. Abbreviations SPA (spatial profile algorithm) and NSPA (non spatial profile algorithm) were used in the following text to denote these two algorithms. NSPA also had a two-layer structure: in the first layer, all the non-target epochs regardless of their *d* to target, were used with target epochs to train the SVM model; in the second layer, a simple classification criterion was applied: trials with all 6 epoch-based p-values below certain threshold will be regarded as non-control, otherwise the trials will be classified to the row/column epoch showing the largest epoch-based p-value.

As there were no 'non-control state' data available from the dataset in use, a simulated dataset with 180 non-control trials was constructed for each subject by randomly concatenating 6 out of all his/her non-target epochs.

To obtain a better signal-to-noise ratio, the epochs used for the algorithms were 3-epoch time averaged epochs from the original data. The first half of data were used to train the SVM models in the algorithms and classification accuracies from the second half of data were reported. A 5×5 fold cross validation procedure was employed to determine the hyperparameters in the SVM models. Matlab 7.5 (The Mathworks, USA), and libSVM [17] were used for data analysis. Linear kernel was adopted when using SVM.

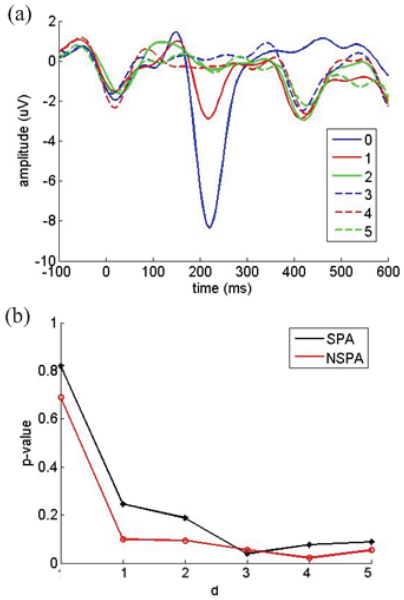

Fig. 2. The spatial profile of the mVEP response: (a) averaged mVEPs from electrode P7 of one representative subject, the waveforms corresponds to different epoch types, according to the *d* values; (b) Transformation of the epoch data into probability outputs.

III. Results

A. Information of the spatial profile in the feature space

Fig. 2 showed the averaged mVEP waveforms of different epochs categorized by the *d* values. For this subject, the N200

component which was the most important feature for classification, declined in amplitude with increasing *d*. A similar trend was also observed from the epoch-based p-values using SPA but not NSPA, indicating the spatial profile was preserved by the first layer classifier.

B. Target classification accuracy

We first calculated the target classification accuracy, defined as the proportion of trials in which the user's intention was correctly recognized, without considering the non-control dataset. Although SPA used a trial (6 epochs) instead of a single epoch for classification, performance comparable to that of NSPA was achieved ($p>0.1$, column ACCURACY, Table I), indicating the effectiveness of SPA in target classification.

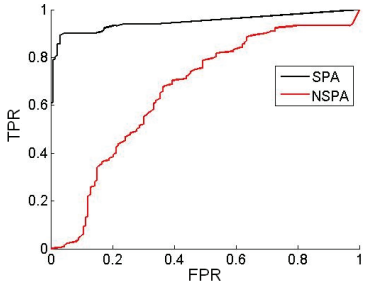

Fig. 3. ROC curve of subject No.8 using SPA and NSPA.

C. Asynchronous performance

To show the general performance of SPA without a subjectively selection of threshold, a receiver operating characteristic (ROC) curve was drawn and the area under the ROC curve was calculated as an index for the performance by sweeping over p-value from 0 to 1. Area of ROC closer to 1 corresponds to better asynchronous performance. Fig. 3 depicted the ROC curve from subject No. 8, with ROC area of 0.60 using NSPA and 0.92 using SPA. Significant better performance was obtained by our proposed algorithm SPA, compared to NSPA ($p<0.0001$, paired t-test).

TABLE I
PERFORMANCE OF THE PROPOSED ALGORITHM(SPA) AND NSPA

Subject	ACCURACY (%)		ROC	
	NSPA	SPA	NSPA	SPA
1	37.8	44.9	0.20	0.35
2	95.5	92.3	0.69	0.88
3	98.1	97.4	0.67	0.97
4	77.0	69.2	0.54	0.57
5	80.8	76.3	0.51	0.62
6	80.1	74.4	0.55	0.64
7	66.7	65.4	0.41	0.58
8	93.0	92.3	0.60	0.92
9	82.7	83.3	0.44	0.74
10	69.2	66.0	0.40	0.56
Mean	78.1	76.2	0.50	0.68

IV. DISCUSSIONS AND CONCLUSIONS

In this paper, a novel algorithm aiming at enhancing the asynchronous performance of the N200 speller was proposed. Utilizing the eccentricity effects of mVEPs, we got a better and more precise description of the brain responses to motion stimuli by integrating the spatial profile of the visual speller. Offline classification showed that 1) the proposed trial-based SPA acquired comparable target classification accuracies with that of the epoch-based NSPA, although a stricter criterion was applied; 2) more importantly, significantly better detection rate of the non-control state was obtained, without using non-control state data for training the classifier.

The idea of doing classification based on trials instead of epochs could be useful for other BCI paradigms such as P300, since the epochs within one trial may correlate with each other. By exploring the relationship between these epochs, better detection of non-control state can be realized.

As a next step, an online system will be implemented to validate the proposed algorithm using real non-control state data. Also, efforts will be devoted to optimizing the layout of the visual speller using SPA algorithm as a validation tool and this will in turn help improve the system performance.

REFERENCES

[1] F. Guo, et al., "A brain-computer interface using motion-onset visual evoked potential," J. Neural Eng., vol. 5, pp. 477-85, Dec 1 2008.

[2] B. Hong, et al., "N200-speller using motion-onset visual response," Clin Neurophysiol, vol. 120, pp. 1658-66, Sep 1 2009.

[3] E. Donchin, et al., "The mental prosthesis: Assessing the speed of a P300-based brain-computer interface," IEEE Transactions on Rehabilitation Engineering, vol. 8, pp. 174-179, Jun 1 2000.

[4] L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," Electroencephalogr Clin Neurophysiol, vol. 70, pp. 510-23, Dec 1988.

[5] M. Kuba and Z. Kubova, "Visual evoked potentials specific for motion onset," Doc Ophthalmol, vol. 80, pp. 83-9, 1992.

[6] W. Skrandies, et al., "Scalp distribution components of brain activity evoked by visual motion stimuli," Exp Brain Res, vol. 122, pp. 62-70, Sep 1998.

[7] M. Korth, et al., "Motion-Evoked pattern visual evoked potentials in glaucoma," J Glaucoma, vol. 9, pp. 376-87, Oct 2000.

[8] E. Dodt and M. Kuba, "Simultaneously recorded retinal and cerebral potentials to windmill stimulation," Doc Ophthalmol, vol. 89, pp. 287-98, 1995.

[9] M. Kuba, et al., "Motion-onset VEPs: characteristics, methods, and diagnostic use," Vision Res, vol. 47, pp. 189-202, Jan 2007.

[10] T. Liu, et al., "An online brain-computer interface using non-flashing visual evoked potentials," J Neural Eng, vol. 7, p. 036003, Jun 1 2010.

[11] J. Kremlacek, et al., "Effect of stimulus localisation on motion-onset VEP," Vision Res, vol. 44, pp. 2989-3000, Dec 2004.

[12] L. Schlykowa, et al., "Motion-onset visual-evoked potentials as a function of retinal eccentricity in man," Brain Res Cogn Brain Res, vol. 1, pp. 169-74, Oct 1993.

[13] A. Lenhardt, et al., "An adaptive P300-based online brain-computer interface," IEEE Trans Neural Syst Rehabil Eng, vol. 16, pp. 121-30, Apr 1 2008.

[14] M. Thulasidas, et al., "Robust classification of EEG signal for brain-computer interface," IEEE Trans Neural Syst Rehabil Eng, vol. 14, pp. 24-9, Mar 1 2006.

[15] H. Zhang, et al., "Asynchronous P300-based brain-computer interfaces: a computational approach with statistical models," IEEE Trans Biomed Eng, vol. 55, pp. 1754-63, Jun 1 2008.

[16] F. Guo, et al., "A brain computer interface based on motion-onset VEPs," Conf Proc IEEE Eng Med Biol Soc, vol. 2008, pp. 4478-81, Jan 1 2008.

[17] C. Chang and C. Lin, "LIBSVM: a library for support vector machines," Software available at http://www. csie. ntu. edu. tw/cjlin/libsvm, 2001.